Contamination in trials of educational interventions

MR Keogh-Brown,[1] MO Bachmann,[1*] L Shepstone,[1] C Hewitt,[2] A Howe,[1] CR Ramsay,[3] F Song,[1] JNV Miles,[4] DJ Torgerson,[2] S Miles,[1] D Elbourne,[5] I Harvey[1] and MJ Campbell[6]

[1] School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK
[2] Department of Health Sciences, University of York, UK
[3] Health Services Research Unit, University of Aberdeen, UK
[4] RAND Corporation, Santa Monica, California, USA
[5] London School of Hygiene and Tropical Medicine, UK
[6] Medical Statistics Group, School of Health and Related Research (ScHARR), The University of Sheffield, UK

* Corresponding author

Executive summary

Health Technology Assessment 2007; Vol. 11: No. 43

Health Technology Assessment
NHS R&D HTA Programme
www.hta.ac.uk

Executive summary

Objectives

The objectives of the study were to consider the effects of contamination on the magnitude and statistical significance (or precision) of the estimated effect of an educational intervention, to investigate the mechanisms of contamination, and to consider how contamination can be avoided.

Background

Educational interventions aimed at improving health, knowledge or health-related behaviour may be delivered to patients, health professionals or members of the general public. Contamination in controlled trials occurs when people who were not intended to receive an intervention inadvertently do so. Trials of educational interventions are especially prone to contamination because the active ingredients can be transportable and difficult to confine. Contamination tends to reduce the magnitude of effect estimates and therefore also to increase the chance that estimates will not be statistically significant. That is, contamination causes bias and reduces power.

Contamination can be avoided during the design, conduct or analysis of trials, but such strategies may be ineffective or may be in conflict with each other. With cluster randomised trials, groups of people are allocated to receive or not to receive an intervention, or to receive different interventions. This reduces contamination bias if it effectively separates people and so reduces the risk or extent of contamination. However, if, within each group, individuals are very similar to each other, compared with individuals in other groups, then the statistical significance and precision of effect estimates are reduced. Various methods of data analysis may adjust for this bias if the extent of contamination is known. However, these adjustments may reduce power and precision because they exclude part of the study sample, or they may cause bias by comparing dissimilar subgroups of the sample. Contamination is often assumed to be a problem when interpreting or designing trials of educational interventions, but whether it really is a problem is not well known. It has been argued that the problems of contamination have been exaggerated, and therefore cluster randomised trials are often inappropriate, given their statistical disadvantages. However, cluster randomisation may be appropriate if interventions are aimed at professionals or facilities that manage groups of people, regardless of contamination.

Methods

An exploratory literature search was conducted with major electronic databases being searched up to May 2005. The results of trials included in previous relevant systematic reviews were then analysed to see whether studies that avoided contamination resulted in larger effect estimates than those that did not. Experts' opinions were elicited about factors more or less likely to lead to contamination. We simulated contamination processes to compare contamination biases between cluster and individually randomised trials. Statistical adjustment was made for contamination using Complier Average Causal Effect analytic methods, using published and simulated data. The bias and power of cluster and individually randomised trials were compared, as were Complier Average Causal Effect, intention-to-treat and per protocol methods of analysis.

Results

Literature search

Although many studies have reported using cluster randomisation to avoid contamination, few have quantified contamination and its effects. We compared the results of cluster and individually randomised trials from previous systematic reviews of educational interventions aimed at health professionals and at patients. We examined whether individually randomised trials tended to show smaller effects, which could indirectly indicate contamination bias. This was true for the relatively few trials that evaluated very similar interventions, although there may be explanations other than contamination or its avoidance. It was not true for the larger number of heterogeneous trials. One interesting trial randomised patients either to a cluster or an individually randomised

sub-trial, both of which evaluated the same oral health intervention. Its results suggested that cluster randomisation reduced contamination bias, but some partial contamination also occurred in the cluster randomised trial. The results could also be explained by unblinded outcome measurement.

Consensus of expert opinion

Thirty-seven experts in trials of educational interventions took part in a Delphi study. They answered a questionnaire ranking the likelihood that contamination would occur in various situations, assuming for each situation that all other factors were constant; 27 completed the second-round questionnaire after feedback on all responses to the first round. In the experts' opinion, contamination was more likely in trials conducted in settings where subjects worked, lived or interacted closely together and where interventions were desirable, simple or easily transferable or were aimed at increasing knowledge. It was less likely when subjects were socially or physically separate, and where interventions were complex or aimed at changing behaviours. It was more likely with interventions aimed at health professionals than with interventions aimed at patients. It was more likely with interventions based on broadcast media, audiovisuals or written information and was least likely with computer-based reminders. Cluster randomisation was the design most likely to avoid contamination and individual randomisation was least likely to do so.

Simulating contamination

A computer model simulated the process of contamination to compare bias between cluster and individually randomised trials. When contamination is not cluster-wide but filters slowly amongst individuals, a cluster randomised design produces less bias. Individual randomisation produces less bias if entire clusters are contaminated at once, unless the risk of clusters being contaminated is low. Different combinations of the components of the contamination process favour either cluster or individual randomisation and should be considered when designing trials. Empirical evidence of the process of contamination during educational trials would be valuable.

Dealing with contamination using Complier Average Casual Effect analysis

Intention-to-treat analysis only answers the question of whether the offer of treatment to the intervention population is effective. Per protocol or on-treatment analyses, which attempt to account for contamination or non-adherence, are likely to be biased because of systematic differences between exposed and unexposed control subjects. Complier Average Causal Effect (CACE) analysis potentially overcomes these problems. It can be implemented in various ways. In this work, we used an instrumental variable technique entailing a two-stage regression. In the first stage, a dummy variable representing the treatment that the participants actually receive is regressed on a dummy variable representing the treatment to which the participants were randomised. Then in the second stage, the outcome is regressed on the treatment received variable and the residuals saved from the first stage. CACE tends to produce an unbiased estimate of the true treatment effect but also tends to reduce statistical power. To assess whether or not the power lost due to CACE analysis was better or worse than that due to cluster randomisation, we undertook a simulation exercise. With up to 30% contamination and using a CACE approach, individual randomisation was more powerful than cluster randomisation. This was true even when assuming small cluster sizes and intracluster correlation coefficients. Although analysis by intention-to-treat is generally the most valid primary analytic method for randomised trials, these methods may be appropriate for secondary analysis of randomised trials, or for analysis of non-randomised trials, if contamination has been measured.

Conclusions

The literature search found little evidence that contamination really is a problem in trials of educational interventions in health because very few studies reported whether contamination occurred. However, there is consensus about the types of situation in which contamination is more or less likely. If it is likely then cluster randomisation may reduce contamination unless entire clusters are contaminated. CACE analysis may reduce bias if contamination is measured. In future trials of educational interventions in health, it is a priority to report the extent, nature and effects of contamination.

Publication

Keogh-Brown MR, Bachmann MO, Shepstone L, Hewitt C, Howe A, Ramsay CR, *et al*. Contamination in trials of educational interventions. *Health Technol Assess* 2007;**11**(43).

NIHR Health Technology Assessment Programme

The Health Technology Assessment (HTA) programme, now part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the costs, effectiveness and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The research findings from the HTA Programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA Programme is needs-led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, the public and consumer groups and professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA Programme then commissions the research by competitive tender.

Secondly, the HTA Programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Thirdly, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer-reviewed by a number of independent expert referees before publication in the widely read monograph series *Health Technology Assessment*.

Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the National Coordinating Centre for Research Methodology (NCCRM), and was formerly transferred to the HTA Programme in April 2007 under the newly established NIHR Methodology Panel. The HTA Programme project number is 06/90/20. The contractual start date was in May 2004. The draft report began editorial review in March 2007 and was accepted for publication in April 2007. The commissioning brief was devised by the NCCRM who specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

ISSN 1366-5278

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.

Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.